U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
SYSTEMS DEVELOPMENT OFFICE
TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 76-12

TESTING LFM-BASED STATISTICAL CLOUD PREDICTION EQUATIONS FOR THE COOL SEASON

Gary M. Carter and Karl F. Hebenstreit*

August 1976

*Lt. Col., USAF, Air Force Liaison Officer assigned to TDL

TESTING LFM-BASED STATISTICAL CLOUD PREDICTION EQUATIONS FOR THE COOL SEASON

by

Gary M. Carter and Karl F. Hebenstreit

The Techniques Development Laboratory (TDL) has been providing automated cloud amount guidance forecasts on teletypewriter for National Weather Service (NWS) forecasters since December 1974. We use single station prediction equations based on the Model Output Statistics (MOS) technique (Klein and Glahn, 1974) to generate probability estimates for categories of opaque sky cover that correspond roughly to clear, scattered, broken, and overcast. These four-category probability forecasts also are transformed into a categorical forecast and presented as a "best" category in the teletype message.

Generally, TDL's objective cloud forecasts compare very favorably with local forecasts prepared at Weather Service Forecast Offices (see Carter et al., 1976). However, field forecasters occasionally notice that these predictions are inconsistent with our objective guidance forecasts of ceiling height.

In an attempt to improve the consistency between automated cloud and ceiling forecasts, we conducted a test involving the simultaneous development of prediction equations for these two weather elements. Since low ceilings occur rather infrequently at some stations (see NWS, 1974a), we divided our 0000 GMT cycle developmental data from the cool seasons (October-March) of 1972-1973, 1973-1974, and 1974-1975 into several forecast regions. Generalized operator equations were then derived for each region by pooling data from all the stations in that region. Fig. 1 shows the 21 regions we selected. These regions were determined subjectively from an analysis at 233 stations of the relative frequency of differing categories of cloud amount and ceiling height when various relative humidity forecasts from the Limited-area Fine Mesh (LFM) model (Howcroft and Desmaris, 1971) were above certain critical values. We also considered topography and climatology wherever possible. Using forecast fields from the LFM model and 0300 GMT observed weather elements, we derived separate equations for each forecast region and projections of 12 and 24 hours from 0000 GMT.

By deriving these "early" guidance (LFM-based) cloud and ceiling prediction equations simultaneously, we were able to insure that all the equations for any given region and projection used the same 12 predictors. This increases the likelihood that the forecasts from these equations will be consistent. However, we questioned if the cloud predictions from the regionalized equations would be as accurate as those from our traditional single station equations.

Using basically the same developmental data, we derived another set of 12- and 24-hour cloud equations for the 40 widely distributed test stations in Fig. 1. Our next step was to comparatively verify cloud forecasts from both systems.

We tested the regionalized and single station cloud prediction equations on independent data from the cool season of 1975-1976. The four-category probability estimates were converted into single best category estimates using an "inflation" technique. Specifically, each probability was adjusted using a statistically determined mean value and correlation coefficient for that particular station or region, category, and forecast projection (see NWS, 1974b). We selected the category with the largest adjusted probability as the categorical forecast. Forecast-observed, four-category contingency tables were then prepared. We computed percents correct, skill scores, and bias-by-category (i.e., the number of forecasts in a particular category divided by the number of observations in that category) for each of the 40 test stations. The combined results are given in Table 1, while Figs. 2 and 3 show the percents correct and skill scores for each station.

The percents correct and skill scores in Table 1 indicate that the regionized equation forecasts are slightly more accurate overall. In contrast, the bias-by-category values favor the single station equation predictions (i.e., each single station bias is closer to unity). Figs. 2 and 3 show that there are primarily only small differences in the percents correct and skill scores from the two systems on a station by station basis for both the 12- and 24-hour forecasts.

In conjunction with this test, we also investigated the merits of using new predictand categories to develop our single station cloud prediction equations. Another set of equations was derived using three "inclusive" predictand categories of clear (1), clear and scattered (2), and clear, scattered, and broken (3). We hoped that these new equations would contain predictors better suited to prediction of the more difficult to forecast scattered and broken categories. The same 40 stations and developmental data were used as in our regional versus single station equation experiment.

We generated test forecasts from both these new inclusive predictand equations and the traditional (single station) exclusive predictand equations from our earlier verification. Once again, the test period was the cool season of 1975-1976, and the two forecast projections were 12 and 24 hours from 0000 GMT.

All the three-category inclusive predictand forecasts had to be converted by a straightforward subtraction technique into the usual four-category probability estimates. Since correlation coefficients were not readily available for these converted categories, the inflation transformation was not used in determining the best categories. Instead, for each system we selected the category with the highest probability as the best category forecast. Although this transformation generally produces forecasts with undesirable bias characteristics, it does offer a means for comparing the overall accuracy between the two forecast systems. Table 2 gives the inclusive and exclusive predictand method verification scores for all 40 stations combined.

The percents correct and skill scores shown in Table 2 are nearly equal for both systems. However, while unacceptable from an operational standpoint, the bias values for the exclusive predictand forecasts are a little closer to unity.

In conclusion, these experiments demonstrate that cloud prediction equations based-on either regionalized or inclusive predictand techniques can produce categorical predictions which are of the same overall accuracy as those produced by our traditional cloud prediction equations. The regionalized approach, which also provides for a greater degree of consistency between TDL's cloud and ceiling forecasts, appears to offer the most encouragement for use in future development of operational equations.

ACKNOWLEDGMENTS

We are grateful to John Neander of MATSCO/GE and Rich Crisci of TDL for preparing the developmental data used in this study. We also appreciate Anna Booth's assistance in drafting the figures shown here.

REFERENCES

Carter, G. M., J. R. Bocchieri, R. L. Crisci, and G. W. Hollenbaugh, 1976: Comparative verification of guidance and local aviation/public weather forecasts--no. 1. TDL Office Note 76-13, Techniques Development Laboratory, Silver Spring, 32 pp.

Howcroft, J., and A. Desmaris, 1971: The limited area fine mesh (LFM) model. NWS Tech. Proc. Bull., No. 67, 11 pp.

Klein, W. H., and H. R. Glahn, 1974: Forecasting local weather by means of model output statistics. Bull. Am. Meteor. Soc., 55, 1217-1227.

National Weather Service, 1974a: The use of model output statistics for predicting ceiling and visibility. NWS Tech. Proc. Bull., No. 120, 10 pp.

______, 1974b: Cloud amount forecasts based on model output statistics (MOS). NWS Tech. Proc. Bull., No. 124, 9 pp.

Figure 1. Twenty-one developmental regions and 40 stations for testing early guidance cloud prediction equations.

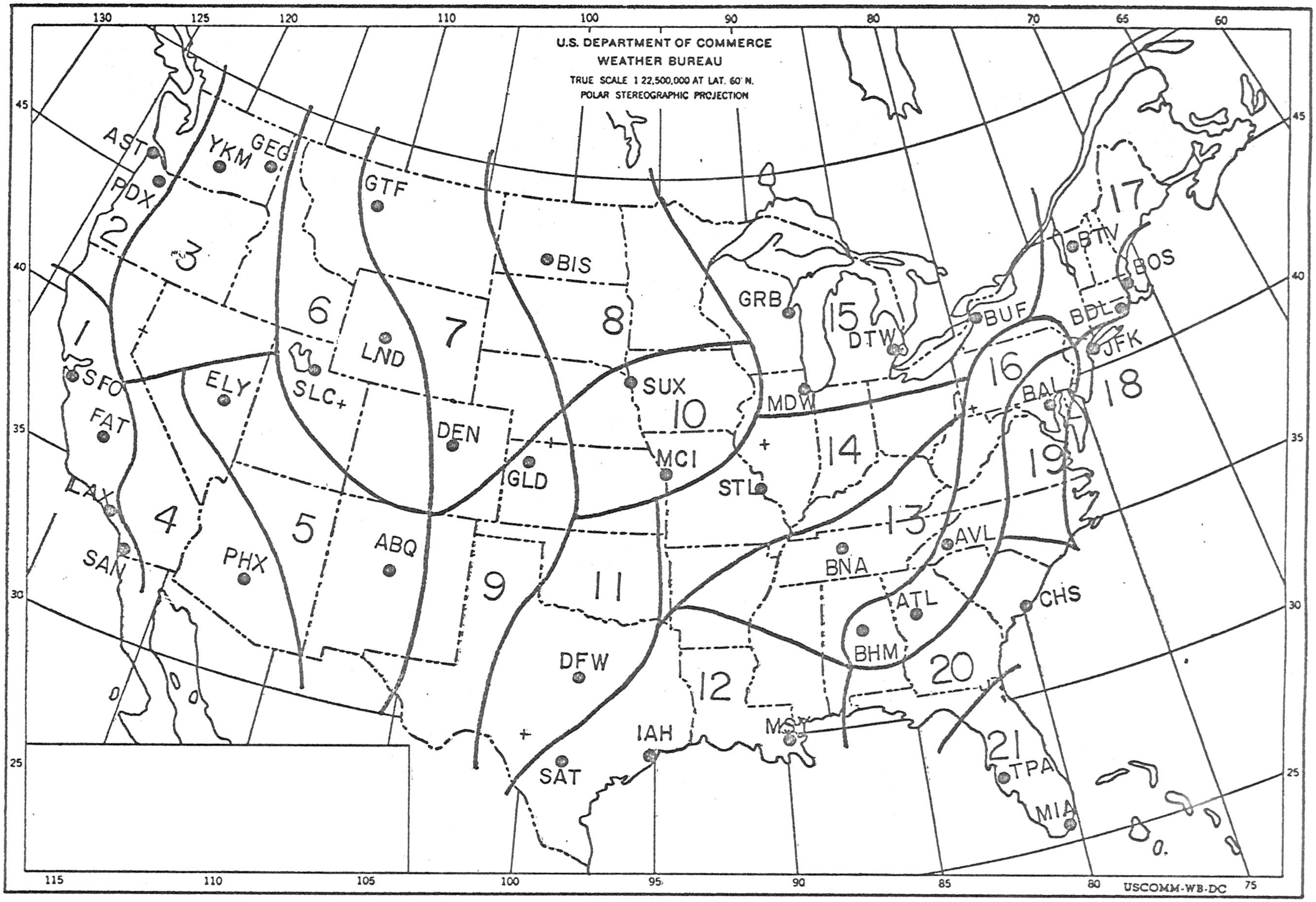

Figure 2. Percents correct and skill scores for 40 U.S. stations during October 1975 through March 1976 for 12-hour predictions of cloud amount based on regionalized and single station techniques.

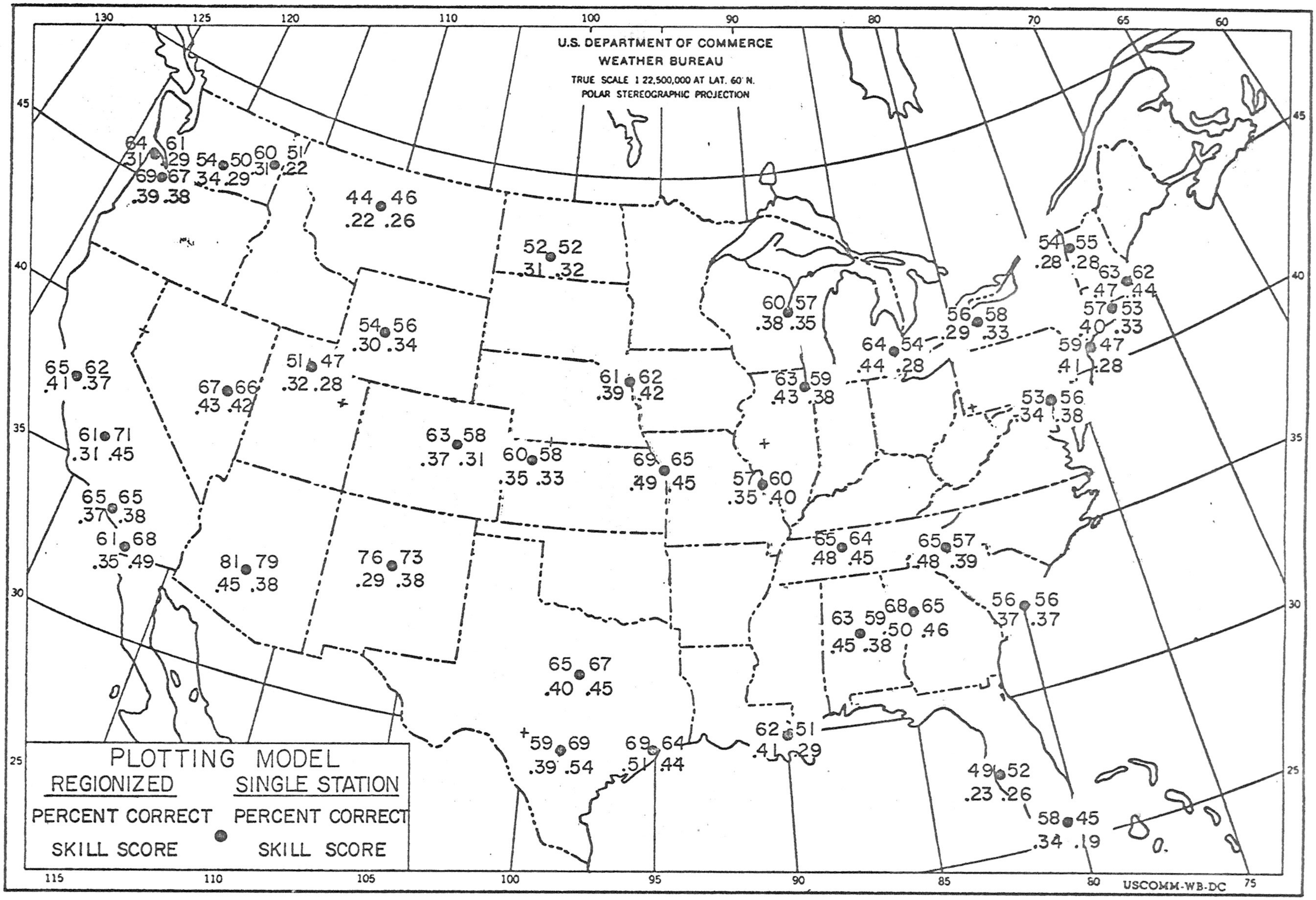

Figure 3. Same as Fig. 2 except for 24-hour predictions of cloud amount.

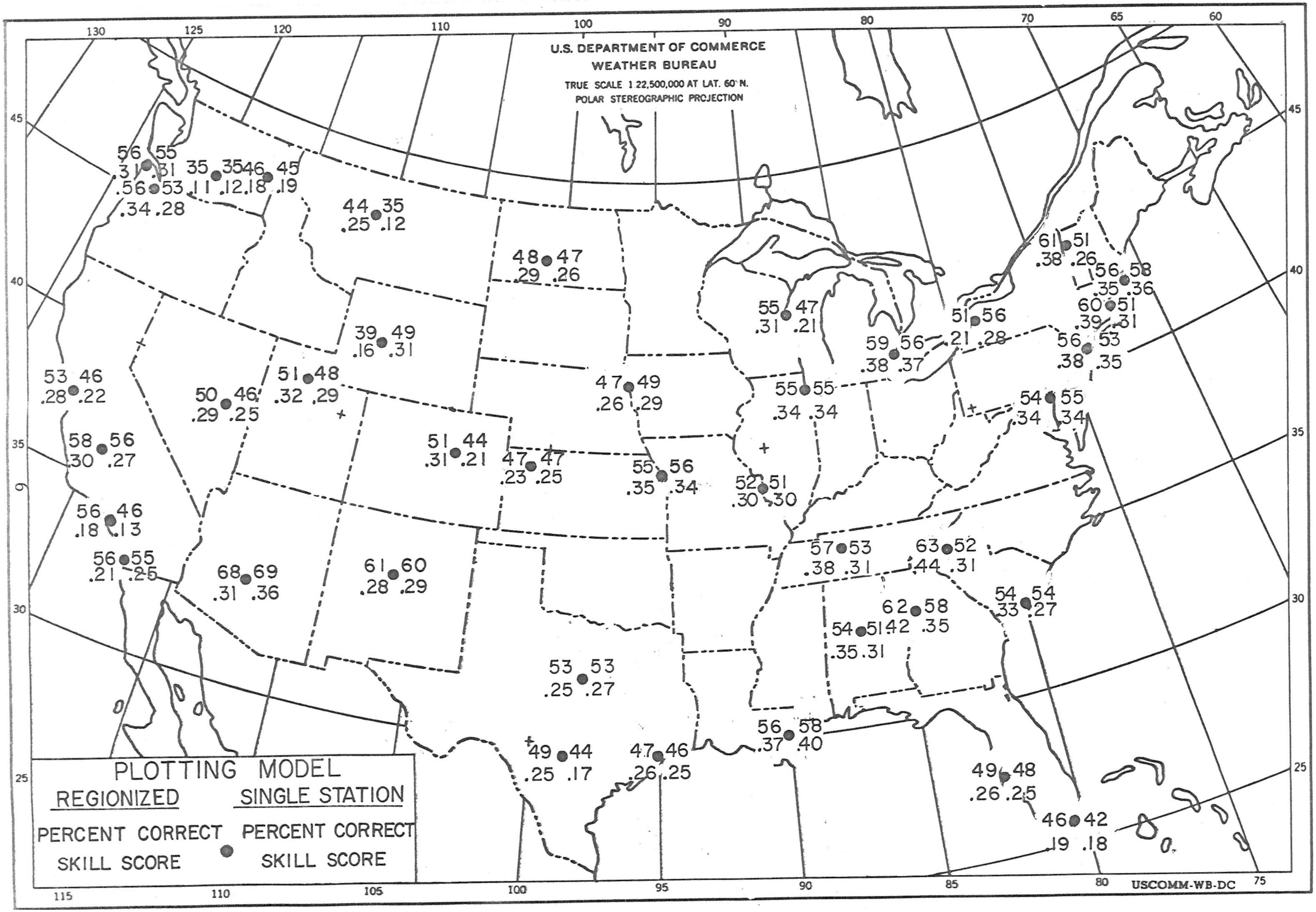

Table 1. Combined verification scores for 40 U.S. stations during October 1975 through March 1976 for predictions of four categories of cloud amount (clear, scattered, broken, and overcast) based on regionalized and single station techniques.

PROJECTION (HOURS)	TYPE OF FORECAST	BIAS - NO. FCST/NO. OBS				PERCENT CORRECT	SKILL SCORE	NO. OF CASES
		CAT 1 (No. Obs.)	CAT 2 (No. Obs.)	CAT 3 (No. Obs)	CAT 4 (No. Obs.)			
12	Regionalized	1.17	0.75	0.60	1.06	61	0.42	6238
	Single Station	1.07	0.85	0.87	1.03	59	0.41	
		(2304)	(1264)	(978)	(1693)			
24	Regionalized	1.10	0.84	0.75	1.13	53	0.35	6239
	Single Station	1.05	0.90	0.92	1.05	51	0.32	
		(2450)	(926)	(769)	(2093)			

Table 2. Same as Table 1 except based on inclusive and exclusive predictand techniques.

PROJECTION (HOURS)	TYPE OF FORECAST	BIAS - NO. FCST/NO. OBS				PERCENT CORRECT	SKILL SCORE	NO. OF CASES
		CAT 1 (No. Obs.)	CAT 2 (No. Obs.)	CAT 3 (No. Obs)	CAT 4 (No. Obs.)			
12	Inclusive	1.23	0.41	0.23	1.27	63	0.43	6238
	Exclusive	1.21	0.45	0.31	1.25	62	0.43	
		(2304)	(1264)	(978)	(1693)			
24	Inclusive	1.26	0.55	0.46	1.30	54	0.35	6239
	Exclusive	1.24	0.57	0.52	1.27	53	0.34	
		(2450)	(926)	(769)	(2093)			